

Tuning of Customer Relationship Management (CRM) via Customer Experience Management (CEM) using Sentiment Analysis on Aspects Level

Hamed AL-Rubaiee

Department of Computer Science and Technology,
University of Bedfordshire
Bedfordshire, United Kingdom

Renxi Qiu

Department of Computer Science and Technology
University of Bedfordshire
Bedfordshire, United Kingdom

Khalid Alomar

Department of Information Systems
King Abdulaziz University
Jeddah, Kingdom of Saudi Arabia

Dayou Li

Department of Computer Science and Technology
University of Bedfordshire
Bedfordshire, United Kingdom

Abstract—This study proposes a framework that combines a supervised machine learning and a semantic orientation approach to tune Customer Relationship Management (CRM) via Customer Experience Management (CEM). The framework extracts data from social media first and then integrates CRM and CEM by tuning and optimising CRM to reflect the needs and expectations of users on social media. In other words, in order to reduce the gap between the users' predicted opinions in CRM and their opinions on social media, the existing data from CEM will be applied to determine the similar behavioural patterns of customers towards similar outcomes within CRM. CRM data and extracted data from social media will be consolidated by the unsupervised data mining method (association). The framework will lead to a quantitative approach to uncover relationships between the extracted data from social media and the CRM data. The results show that changing some aspects of the e-learning criteria that were required by students in their social media posts can help to enhance the classification accuracy in the learning management system (LMS) data and to understand more students' studying statuses. Furthermore, the results show matching between students' opinions in CRM and CEM, especially in the negative and neutral classes.

Keywords—Opinion mining; customer relationship management; customer experience management; sentiment analysis; Twitter

I. INTRODUCTION

Today, social media has become a method for maintaining strong relationships between users and companies.¹ With social media, companies are no longer in control of the relationship. Instead, customers are now driving the conversation [1]. To meet users' expectations and understand their opinions via social media platforms, companies are keen to adjust their CRM based on the feedback from social media platforms. In other words, CRM can be fine-tuned based on the differences between its own prediction of users' opinions and the actual

feedback from CEM. The differences between the opinions predicted by CRM and the actual opinions collected from social media are important to improve CRM further. The inconsistency in the opinions comes from two sources of error in CRM, the criteria or weights employed in CRM are not accurate, and the structure of CRM is not optimised.

In the former case, CRM and CEM cover similar aspects in a given domain; their difference is solely due to the inaccuracy of the criteria or weights employed in the CRM system. In the latter case, CRM could have missed a few key aspects that are important for CEM. The difference in the latter case is inherent in the structure of the CRM system.

Facing the problem, tuning CRM via CEM needs to be organised in two levels:

- Tuning on the aspect level: In this case, CRM and CEM have been constructed from similar aspects; sentiment analysis could be applied to adjust the criteria or weights in the CRM via the difference between CRM and CEM. This part of the problem will be investigated in this paper. A framework for adjusting CRM via CEM will be proposed and validated by a real-word example from the educational sector.
- Tuning on the sentence level: In this case, CRM and CEM have been constructed separately with different aspects. Opinions can only be collected on the sentence level for certain subjects in this domain.

This paper organised as follows: In Section 2, a related work for integrating CRM and CEM is introduced. In Section 3, a framework for integrating CRM and CEM is introduced. Section 4 contains the CRM tuning based on CEM at the aspect level—a case study from King Abdul-Aziz University. Section 5 presents sentiment analysis for CEM along with experimental results and evaluations. In Section 6, classification using CRM along with results and evaluations is

¹ <https://rapidminer.com>

shown. In Section 7, tuning of CRM via CEM is presented along with experimental results and evaluations. Finally, a summary of this paper is presented in Section 8.

II. RELATED WORK

In the first instance, a sentiment analysis system can be developed to determine consumer attitudes on products/services from review data. For example, when Dehkharghani and Yilmaz [2] conducted a review using a logistical classifier, they found that the average accuracy was 66.6% [3]. Bross and Ehrig [4] found that an aspect-based review to detect individual opinions and expressions about specific aspects of a product had a high accuracy [3]. Indhuja and Reghu [5] used an approach with novel fuzzy functions and achieved 85.58% accuracy. Wang [6] found that his model using a combined sentiment LDA and topic LDA was more effective than just topic sentiment analysis. Zhang and Varadarajan's [7] models, which incorporated features to predict utility scores of product reviews, achieved high performance, indicating that this is an effective approach [3].

The second example looks at the influence of Twitter; in terms of any relationship between sentiment analysis and the stock market, Bollen, Mao, and Zeng [8] reported that Twitter moods were used to predict the Dow Jones stock market index [9]. In their approach, public moods were measured using two tools: Google Profile of Mood (GPOM) and Opinion Finder [9] and a predictive model based on Self-Organised Fuzzy Neural Networks (SOFNN). The study found that the accuracy of the standard stock prediction model was significantly improved when mood dimensions and the value of the Dow Jones Industrial Average (DJIA) were included [9]. In addition, a study by Martin, Bruno, and Murisasco [10] showed a correlation between public opinion expressed on Twitter and the French stock market, using a neural network to find association patterns. They found that adding the sentiment feature on tweets two days before stock market closing values could improve accuracy [9].

Simsek and Ozdemir [11] also found a relationship between Turkish Twitter posts and the stock market index. By using the most common Turkish words representing happiness and unhappiness collected from tweets over a period of a month and a half, they worked out the frequencies of these two classes of words [9]. They found that the terms happy and trouble were commonly used in the emotional word database, and when the Twitter post contained stock market-related words, the average emotional value of the tweets changed from happiness to unhappiness by approximately 45% [9]. In a further study, Khatri, Singhal, and Johri [12] tried to train the neural network with words such as happy, hope, sad, and disappointing as input to predict the Bombay Exchange index. Their study indicated that an artificial neural network provided optimum results when set up with one hidden layer containing nine neurons [9]. Gao et al. [13] tried using a sentiment classification approach for Chinese stock news and found that pre-processing and a relevant sentiment dictionary affected the classification. Zhang and Skiena [14] showed that news sentiment can have an influence on market trading algorithms; they also found that news sentiment had a much sooner impact on stock markets than sentiment expressed on Twitter, which

could take up to three days [9]. All of these studies, however, point to a strong correlation between sentiments expressed on Twitter and global stock market indices.

In the third example, student feedback can highlight any issues students may have with the services provided by their colleges or universities. An example of this is when students cannot understand a lecturer or do not avail themselves of specific online services. Students have a habit of regularly using social media to express their opinions and describe activities they are involved in [15]. Therefore, universities utilise social media as a way of improving their teaching processes and generally to find out more about student experiences [15]. This is especially useful when finding out about online distance-education students, who do not give feedback face to face [15]. For instance, Tian et al. [16] developed an e-learner questionnaire and compared emotion words to measure the intensity of sentiment in each category. This approach enabled a positive result in dealing with challenges faced in the analysis of texts, such as those in Chinese, which are characterised by the richness of emotions. Wang, Zuo, and Diao [17] worked with the essential function of sentiment feedback in education over the Internet [15]. Wang [18] set up a Student Feedback Mining System (SFMS) to carry out an in-depth analysis of qualitative student feedback, which allowed insight into teaching practices, thus significantly improving student learning [15]. Donovan, Mader, and Shinsky [19] found that online student comments were much more detailed and informative than traditional paper-based feedback but are more time-consuming to analyse [15]. However, sentiment classification is important, as it gathers attitudes and opinions of users by mining and analysing personal information [18].

In general, a social CRM conversation between customers and enterprise agents over social media is called social CRM. Social CRM can influence the customer community to solve a customer's problem and turn a bad opinion into a good one [20]. Furthermore, social media has been used extensively by enterprises in the recent past to get insights about what users think about their products or services; this is typically achieved in a "listening" mode, i.e., a large amount of data from multiple social media sites is analysed in offline mode to extract aggregate-level business insights [21], [22]. Usually, the relational model database is the backbone for most companies or organizations; it stores the data in a structured order with rows and columns. However, a huge percentage of data in organizations or companies are located in unstructured data such as text data. This shows the need to analyse the unstructured data for companies' benefit [23].

Many researchers have shared their experiences with this subject. Ajmera et al. [24] built a social CRM that enabled firms to engage with customers by presenting analytical methods to identify actionable posts and analysing them. They presented novel features such as user intent and severity of issues in a customer complaint to determine a post's priority. In addition, Yaakub and Zhang [25] proposed a multi-dimensional model for opinion mining to integrate customers' characteristics and their related opinions about products. They used POS tagging to pre-process their data, trying to capture three parts from each document: nouns, which describe the

name of the product, and adverbs and adjectives, which describe the sentiment toward that product. However, it was hard to evaluate their model's conclusions about product attributes based on customers' opinions. In other words, it is hard to cover all details in CRM using only sentiment analysis. The baseline models for the above study were developed by trawling the reviews before putting them into a review database [26].

III. FRAMEWORK OF TUNING CRM VIA CEM

This study proposes a framework that combines a supervised machine learning and a semantic orientation approach to tune CRM via CEM. The framework extracts data from social media first and then integrates CRM and CEM by tuning and optimising CRM to reflect the needs and expectations of users on social media. In other words, in order to reduce the gap between the users' predicted opinions in CRM and their opinions on social media, the existing data from CEM will be applied to determine the similar behavioural patterns of customers towards similar outcomes within CRM. CRM data and extracted data from social media will be consolidated by the unsupervised data mining method (association). The framework will lead to a quantitative approach to uncover relationships between the extracted data from social media and the CRM data.

Fig. 1 illustrates the proposed framework for integrating CRM with CEM. The framework consists of three processes: (1) sentiment analysis for CEM, (2) classification using CRM, and (3) data tuning of CRM via CEM. In terms of data modelling, the main components of the three processes are very similar. The difference between process one and process two is in the input: one takes unstructured social media input, while the other takes a structured customer database. In process three, CRM data can be labelled automatically by CEM or vice versa. In this work, we will focus on the fine-tuning of CRM through CEM.

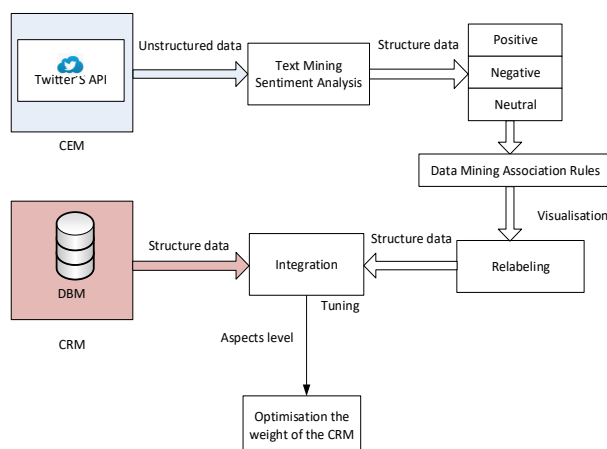


Fig. 1. The process of the proposed framework.

A. Sentiment Analysis for CEM

This is a process that extracts key features for CEM using social media. The extracted features will be represented by a semantic schema. The schema can be applied directly to the social media input. The process:

- Collects tweets based on a set of keywords that describe the case study using Twitter's streaming API.
- Performs text-processing techniques based on the proposed ontology model to reduce the amount of noise.
- Extracts key features with supervised learning. By exploring relevant data mining approaches such as sentiment analysis and NLP on building models, the component classifies tweets according to their sentiment polarity into one of the classes of positive, negative, and neutral.

B. Classification using CRM

This is a process that extracts key features for CRM using an existing customer database. The extracted features are also represented by a semantic schema, which can be applied directly to the consumer database.

- Collects structured data for the CRM using the database's API.
- Performs pre-processing techniques based on the existing CRM model to reduce the amount of noise.
- Extracts key features with supervised learning, which is again similar to the CEM process.

C. Tuning of CRM via CEM

A model has been developed that cross-validates features extracted from both CEM and CRM. In this case, if CRM and CEM are not consistent, CRM's semantic schema can be updated by CEM's output directly. In this process:

- Statistical algorithms will be applied to discover patterns and correlations in features extracted from both CEM and CRM.
- The confidence of the discovery will be examined automatically by comparing validation between the outputs. False positives, negatives, and neutrals will be identified during the process.
- CRM's semantic schema will be revised iteratively. If CRM and CEM are constructed from similar aspects, the tuning will be focused on the criteria or weights in CRM. Otherwise, CRM and CEM will need to be interpreted on the sentence level, so the structure of the CRM will be optimised based on its difference from CEM at this level.

IV. CASE STUDY FOR TUNING CRM VIA CEM AT THE ASPECT LEVEL: OPTIMISATION OF CRM OF KING ABDUL-AZIZ UNIVERSITY

The Deanship of e-Learning and Distance Education at King Abdul-Aziz University's Twitter account was chosen as a platform for opinion mining. The investigation was conducted on students' sentiments on the aspect level, with students being asked specific questions about e-learning criteria. The aim of this case study was to validate the proposed framework by adjusting the criteria or weights employed in the targeted CRM system.

A. Preparing the Sentiment Analysis for CEM using the Criteria from CRM

In order to prepare the sentiment and to clean the text to hand it off to the classifier, links, URLs, and hashtags were removed from the sentiments. In the following case study, the experimental collection process was based on certain hashtags that represent important topics, using Twitter hashtags as the domain.

B. CEM based on Sentiment on the Aspect Level

In this experiment, a hashtag was created to ask users for opinions on and reactions toward distance education criteria in order to identify students' positive, negative, and neutral opinions. The aspect level of sentiment analysis was followed for the students' opinions on the Twitter platform. The collection process was based on certain hashtags that represent the criteria, such as “#طور آلية التعليم عن بُعد” “# Enhance Distance education criteria”. The experiment aimed to illustrate the relationship between the sentiments conveyed in Arabic tweets and the students' learning experiences at universities.

C. Experiment on Sentiment Analysis for CEM

During this stage, a hashtag was created to ask users for opinions on and reactions toward distance education criteria in order to identify students' positive, negative, and neutral opinions.

D. Experimental Parameter Settings

The classifier settings followed the default values for most of the parameters to get more accurate results. Ten-fold cross-validation was used several times to find the best values of these parameters. In addition, both SVM and NB had the same operator parameter settings used.

E. Data Collection and Description

To collect opinions that were comprehensive for the time period on the targeted objectives, tweets from different students on the Deanship of e-Learning and Distance Education page were obtained using Twitter's official developer's API. The data were collected over four months between 14 April and 13 August 2017. The data distribution depended on the methods Twitter's API utilised and the number of tweets posted on the distance education Twitter account page. Downloaded tweets were marked manually by employees of the distance education deanship as positive, negative, or neutral. In addition, the labelled tweets were stored in a database for experiments.

Distance education students' CRM data were obtained from the Learning Management System (LMS) Blackboard at King Abdul-Aziz University, which consists of the most important fields that can describe the students' activities. In this experiment, the collection process was based on students' data from one course chosen randomly from the student schedule of the second semester in 2017. Distance education students were labelled as “neutral” who had total marks in a course between 60 and 75 and attended 9 to 10 out of 12 lectures. For students who had total marks in a course less than 60 and attended fewer than 9 out of 12 lectures, the label was “negative”. If a student had total marks above 75 and attended more than 10 out of 12 lectures, then he or she was considered an active

student and labelled as “positive”. Table I shows the number of tweets that were involved in the classification experiment. Details of 143 students' were collected from the Blackboard CRM. All of them were anonymised. After that the data were labelled depending on distance education criteria, and then the data were stored. Rapidminer was used to load the dataset in normalized form with no duplicate records and no null attribute values.

After the data collection stage, the data model layout was created. This included a tweets description table (Tweet ID, Tweet Original, Tweet Filtered, Tweet Time, Tweet User, Tweet Label) and selected attributes from Blackboard databases. Again, this information was anonymised to protect privacy. Each student record contained the individual student's positive, negative, and neutral opinions, as well as the original tweet and the tweet labelled as an opinion from the CEM data for each individual student. Each student's record thus had two labels. The CRM label depended on the CRM criteria and the CEM label, which came from the sentiment analysis model.

Ten thousand six hundred and three students viewed the hashtag and agreed to start the experiment. Only 567 students registered their Twitter accounts and allowed the university to follow their tweets. One of the obstacles to letting students continue the survey was that the authentication from Twitter is in English. However, 242 of the 567 students have records in Blackboard, since only distance education students must use Blackboard regularly, whereas regular students and external students in some colleges still do not use Blackboard. Of these 242 students, 143 completed the survey.

The similar proposed model for sentiment analysis of Saudi Arabic (Standard and Arabian Gulf dialects) tweets was applied to extract feedback features from King Abdul-Aziz University data. The main idea was to examine the aspect level in sentiment analysis. In addition, the neutral class is important in the Arabic sentiment analyses, so our experiment was carried out with the neutral class as well as the negative and positive classes. The following tables show the type of tuning to provide the best accuracy in sentiment analysis for Arabic text in relation to King Abdul-Aziz University. For this experiment, 143 students' tweets responding to four questions were utilised. Table I shows the total number of tweets for the 143 students utilized for this experiment.

TABLE. I. NUMBER OF TWEETS UTILIZED FOR DATA MODEL COMPONENT ONE

Question	Positive	Negative	Neutral	Total
Q1	73	11	59	143
Q2	54	80	9	143
Q3	87	37	19	143
Q4	39	42	62	143
Total	338	369	261	572

F. Data Pre-Processing

A selection of 28 words and idioms in Arabic from the emotion corpus such as growth, good, excellent, problem, and inappropriate (see Table II) was then formed into the following three classes: positive, negative, and neutral. The most common Arabic words in the standard Saudi dialect represent positive and negative classes.

TABLE. II. COMMON ARABIC WORDS AND STANDARD SAUDI DIALECT REPRESENTING THE CLASSES

Arabic Positive Sentiment	English Translation	Arabic Negative Sentiment	English Translation
فتح تخصصات جديد	open new courses	مشكلة	problem
مناسب	suitable	تطوير بلاك بورد	develop Blackboard
جيد	good	لا يوجد تجاوب	no response
متكامل	complementary	غير مناسب	unsuitable
ممتاز	excellent	غير ملائم	inappropriate
سريع جدا	very fast	غير كافي	inadequate,
نظام جيدة	good platform	حرمان	absence exceeds limits
موافق	agree	زيادة النسبة	Increase percentage

Two distance education employees who have experience in e-learning labelled the data manually. Positive tweets were given the label "1", while negative tweets were given the label "-1". Neutral tweets were given the label "0", and irrelevant tweets were deleted from the database. A survey with four questions for e-learning criteria 2 was created and made available on Twitter via hashtags (see Table III). Three questions asked a specific question on the criteria, while the last one was an open question asking students about their opinions in general about e-learning. Table 4 shows the four questions asked of participants.

After data labelling was completed, the data were stored in our system in normalized form, as mentioned in the framework, with no hashtags, no duplicate tweets, no retweets, no URLs, and no special characters. After loading the dataset, the data were pre-processed with Rapidminer. The first step was to replace some Arabic words taking different shapes and icons.

TABLE. III. E-LEARNING CRITERIA

Number	E-learning criteria
1	Classwork: includes interaction with the instructor through the given activities available in the learning management system (30 marks of the grand total). Final examination: 70 marks of the grand total.
2	If student absence exceeded three lectures (equivalent to 25% of the synchronous online lectures) offered throughout the semester, then the student is prevented from taking the final exam of the distance learning course.
3	The quarterly work is divided as follows: assignments 4 (3 marks per assignment), activities 2 (6 marks per activity), forums (discussion board) 3 (2 marks per forum).
4	Any comments about the distance learning mechanism.

TABLE. IV. A SURVEY WITH FOUR QUESTIONS FOR E-LEARNING CRITERIA

Question No.	Question
Q1	What do you think about the following criteria: The student evaluates from 100 degrees: the quarterly work (30 degrees of the total) and the final test (70 degrees of the total).
Q2	What do you think about the following criteria: The student is denied entry to the final exam of the distance learning course if his absence exceeds 3 (25%) of the lectures.
Q3	What do you think about the following criteria: The quarterly work is divided as follows: four duties for each subject and the calculation of a single assignment three degrees, two periodic tests and each test six degrees, participation in the forums (discussion board) three posts, and each class two degrees (2).
Q4	Are there any comments you would like to share with us about the distance learning mechanism?

Then, the same pre-processing steps were performed: tokenization, removal of stop words, light stemming, filtering tokens by length, and application of the N-gram feature. Next, the 'Process Documents from Data' operator generated a word vector from the dataset after pre-processing and represented the text data as a matrix to show the frequency of occurrence of each term; then, relevant data mining approaches, i.e., NB and SVM, were explored for building models to classify tweets according to their sentiment polarity into positive, negative, and neutral. Finally, an evaluation was carried out using precision and recall methods.

G. Experiment Results and Evaluations

The results were divided into four groups for each question to show the sentiment analysis classification accuracy, precision, and recall for the NB and SVM classifiers with and without the N-gram feature.

• Sentiment analysis classification for Question 1

Table V shows Q1 classification accuracy, precision, and recall for the NB and SVM classifiers without the N-gram feature: Crosse-validation=10, sampling type=stratified sampling, prune=none. In addition, Table VI shows the class accuracy, precision, and recall for the NB and SVM classifiers with the N-gram feature, which is set to two: Crosse-validation=10, sampling type=stratified sampling, prune=none.

In conclusion, the experiment shows that NB performance was better when we used the N-gram feature with both schemas (TF-IDF and BTO). On the other hand, there was a slight performance increase when SVM used the same feature. However, the best accuracy was achieved by SVM with the TF-IDF schema when the N-gram feature was not involved.

²http://elearning.kau.edu.sa/Content.aspx?Site_ID=214&lng=EN&cid=24144

TABLE. V. ACCURACY, PRECISION, AND RECALL FOR ALL CLASSES USING SVM AND NB CLASSIFIERS WITHOUT N-GRAM FOR QUESTION 1

Classes	Classifier Name	Weighting schemes	Accuracy	Class recall	Class precision	Classification error
All Class	NB	BTO	39.24	43.95	48.18	60.76
		TF-IDF	55.19	54.59	54.54	44.81
	SVM	BTO	77.57	55.06	52.73	22.43
		TF-IDF	83.19	59.65	56.74	16.81

TABLE. VI. ACCURACY, PRECISION, AND RECALL FOR ALL CLASSES USING SVM AND NB CLASSIFIERS WITH N-GRAM FOR QUESTION 1

Classes	Classifier Name	Weighting schemes	Accuracy	Class recall	Class precision	Classification error
All Class	NB	BTO	55.24	53.36	54.42	44.76
		TF-IDF	67.86	61.65	62.66	32.14
	SVM	BTO	78.95	61.73	60.42	21.05
		TF-IDF	81.05	63.59	61.97	18.95

TABLE. VII. ACCURACY, PRECISION, AND RECALL FOR ALL CLASSES USING FOR SVM AND NB CLASSIFIERS WITHOUT N-GRAM FOR QUESTION 2

Classes	Classifier Name	Weighting schemes	Accuracy	Class recall	Class precision	Classification error
All Class	NB	BTO	77.62	66.44	63.18	22.38
		TF-IDF	76.24	65.47	62.26	23.76
	SVM	BTO	78.29	56.11	52.81	21.71
		TF-IDF	72.76	49.67	50.48	27.24

TABLE. VIII. ACCURACY, PRECISION, AND RECALL FOR ALL CLASSES USING FOR SVM AND NB CLASSIFIERS WITH N-GRAM FOR QUESTION 2

Classes	Classifier Name	Weighting schemes	Accuracy	Class recall	Class precision	Classification error
All Class	NB	BTO	78.33	66.86	63.54	21.67
		TF-IDF	77.67	66.31	63.06	22.33
	SVM	BTO	74.90	54.53	51.81	25.10
		TF-IDF	70.00	46.94	49.80	30.00

TABLE. IX. ACCURACY, PRECISION AND RECALL FOR ALL CLASSES USING FOR SVM AND NB CLASSIFIERS WITHOUT N-GRAM FOR QUESTION 3

Classes	Classifier Name	Weighting schemes	Accuracy	class recall	Class precision	Classification error
All Class	NB	BTO	72.05	64.58	66.34	27.95
		TF-IDF	76.95	67.22	67.20	23.05
	SVM	BTO	80.33	61.85	62.91	19.67
		TF-IDF	82.52	65.19	68.74	17.48

TABLE. X. ACCURACY, PRECISION AND RECALL FOR ALL CLASSES USING FOR SVM AND NB CLASSIFIERS WITH N-GRAM FOR QUESTION 3

classes	Classifier Name	Weighting schemes	Accuracy	class recall	Class precision	Classification error
All Class	NB	BTO	80.48	70.46	67.66	19.52
		TF-IDF	81.86	71.76	71.15	18.14
	SVM	BTO	79.67	62.69	63.60	20.33
		TF-IDF	81.05	61.39	61.32	18.95

TABLE. XI. ACCURACY, PRECISION, AND RECALL FOR ALL CLASSES USING FOR SVM AND NB CLASSIFIERS WITHOUT N-GRAM FOR QUESTION 4

Classes	Classifier Name	Weighting schemes	Accuracy	Class recall	Class precision	Classification error
All Class	NB	BTO	60.00	61.51	59.15	40.00
		TF-IDF	69.10	71.34	70.53	30.90
	SVM	BTO	58.00	53.27	65.65	42.00
		TF-IDF	68.43	67.48	74.74	31.57

TABLE. XII. ACCURACY, PRECISION, AND RECALL FOR ALL CLASSES USING FOR SVM AND NB CLASSIFIERS WITH N-GRAM FOR QUESTION 4

Classes	Classifier Name	Weighting schemes	Accuracy	Class recall	Class precision	Classification error
All Class	NB	BTO	59.29	61.03	58.54	40.71
		TF-IDF	69.81	71.90	71.59	30.19
	SVM	BTO	49.81	42.33	48.96	50.19
		TF-IDF	62.86	62.89	72.66	37.14

- Sentiment Analysis Classification for Question 2

Table VII shows Q2 classification accuracy, precision, and recall for the NB and SVM classifiers without the N-gram feature: Crosse-validation=10, sampling type=stratified sampling, prune=none. In addition, Table VIII shows the class accuracy, precision, and recall for the NB and SVM classifiers with the N-gram feature, which is set to two: Crosse-validation=10, sampling type=stratified sampling, prune=none.

In conclusion, the experiment shows that NB performance was better when we used the N-gram feature with both schemas (TF-IDF and BTO). On the other hand, there was a performance decrease when SVM used the same feature. However, the best accuracy was achieved by SVM with the BTO schema when the N-gram feature was not involved.

- Sentiment Analysis Classification for Question 3

Table 9 shows Q3 classification accuracy, precision, and recall for the NB and SVM classifiers without the N-gram feature: Crosse-validation=10, sampling type=stratified sampling, prune=none. In addition,

X shows the class accuracy, precision, and recall for the NB and SVM classifiers with the N-gram feature, which is set to two: Crosse-validation=10, sampling type=stratified sampling, prune=none.

In conclusion, the experiment shows that NB performance was better when we used the N-gram feature with both schemas (TF-IDF and BTO). On the other hand, there was a performance drop when SVM used the N-gram feature with both schemas (TF-IDF and BTO).

- Sentiment Analysis Classification for Question 4

Table XI shows Q6 classification accuracy, precision, and recall for the NB and SVM classifiers without the N-gram feature: Crosse-validation=10, sampling type=stratified sampling, prune=none. In addition, Table XII shows the class accuracy, precision, and recall for the NB and SVM classifiers with the N-Gram feature, which is set to two: Crosse-validation=10, sampling type=stratified sampling, prune=none.

In conclusion, the experiment shows that NB performance was better when we used the N-gram feature with both schemas (TF-IDF and BTO). On the other hand, there was a drop when SVM used the N-gram feature with both schemas (TF-IDF and BTO).

V. CLASSIFICATION USING CRM

Component two presented the design and implementation of students' CRM data (or classification through different algorithms, such as SVM and NB. In this part of the experiment, the collection process was based on students' Blackboard attributes assigned according to their points of view on different classes in Blackboard data. It was labelled according to the university's e-learning criteria.

A. Details of using the Classifiers with Rapidminer for CRM Classification

The classifier settings followed the default values for most of the parameters to get more accurate results. Ten-fold cross-

validation was used several times to find the best values of these parameters.

B. Support Vector Machine (SVM) Operator Parameter Settings

The SVM operator generates the SVM classification model. This model can be used for classification and provides good results for many learning tasks. In addition, it supports various kernel types, including dot, radial, polynomial, and neural³.

- SVM type: C-SVM, which is for classification tasks.
- Linear: A linear classifier works based on the value of a linear combination.
- C is the penalty parameter of the error term and was set to its default value, which is zero.
- Cache size is an expert parameter. It specifies the cache size in megabytes and was set to default value, which is 80.
- Epsilon: This parameter specifies the tolerance of the termination criterion and was set to the default value, which is 0.001.

C. Naïve Bayes (NB) Operator Parameter Settings

The NB operator generates a NB classification model. It is a probabilistic classifier based on applying Bayes' theorem with powerful independence assumptions. The NB classifier assumes that the presence or absence of a particular feature of a class is unrelated to any other feature.

- Laplace correction:

This parameter indicates whether Laplace correction should be used to prevent high influence of zero probabilities. Assume that our training set is so large that adding one to each count will make a small difference in the estimated likelihoods.

VI. EXPERIMENT OF THE CLASSIFICATION USING CRM

For this experiment 567 student registered their twitter account and allow the university to follow their tweets. However, only 234 students are distance education students and have full record in Blackboard, where other students do not use Blackboard. Table XIII shows the total number of tweets for the 242 students utilised for this experiment.

TABLE. XIII. NUMBER OF TWEETS UTILIZED FOR DATA MODEL COMPONENT TWO

Question	Positive	Negative	Neutral	Total
Q1	84	65	93	242
Q2	90	122	30	242
Q3	84	89	69	242
Q4	80	93	69	242
Total	338	369	261	968

A. Data Description for Question 1, 2, 3, and 4

Table XIV shows the criteria utilised to label record as positive, negative, or neutral for question 1.

³ <https://rapidminer.com>

TABLE. XIV. QUESTION 1 CRITERIA

	Labelled	Criteria
Student Mark at Blackboard (total mark is 30)	Positive	>=20
	Neutral	15-20
	Negative	<15
Student Final Exam Mark in ODUS from (total mark is 70)	Negative	<35
	Neutral	35-45
	Positive	>45
Student Final Mark in ODUS from (total mark is 100)	Negative	<50
	Neutral	50-75
	Positive	>75
TOT_MARK = BB_F_MARK_NO + ODS_MARK		

Table XV shows the criteria utilised to label record as positive, negative, or neutral for question 1.

TABLE. XV. QUESTION 2 CRITERIA

	Labelled	Criteria
Number of recorded attendance	Negative	>5 days
	Neutral	4-5 days
	Positive	<4 days
Number of online attendance	Negative	<9 lectures
	Neutral	9-10 lectures
	Positive	>10 lectures
ATTEND_MARK = REC_MARK_COUNT + ONLINE_MARK		

Table XVI shows the criteria utilised to label record as positive, negative, or neutral for question 1.

TABLE. XVI. QUESTION 3 CRITERIA

	Labelled	Criteria
Number of times student participate in the discussion, quizzes and assignments	Negative	>5
	Neutral	5-6
	Positive	>6
FORUM_TEST_MARK_TOTAL = FORUM_TEST_MARK_FORUM + FORUM_TEST_MARK_TEST		

Table XVII shows the criteria utilised to label record as positive, negative, or neutral for question 1.

TABLE. XVII. QUESTION 4 CRITERIA

	Labelled	Criteria
Total Mark	Negative	<3
	Neutral	<5
	Positive	>6
TOTAL_EVA_CHAR = (TOT_MARK + ATTEND_MARK + FORUM_TEST_MARK_RESULT) / 3		

VII. EXPERIMENTAL RESULTS AND EVALUATIONS

The results were divided into groups for each question to show the Blackboard students' data classification accuracy, precision, and recall for the NB and SVM classifiers.

A. CRM Classification for Question 1, 2, 3, and 4

The following table shows the type of tuning to provide the best accuracy in King Abdul-Aziz University's CRM classification. Table XVIII shows the classification accuracy, precision, and recall for the NB and SVM classifiers: Crosse-validation=10, sampling type=stratified sampling. The best accuracy was achieved by NB due to the advantages of NB, such as its simplicity, ease of implementation, and combination of efficiency with acceptable accuracy.

TABLE. XVIII. ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS

Classes	Classifier name	Accuracy	Class recall	Class precision	Classification error
Positive, negative, and neutral	NB	93.43	93.95	94.12	6.57%
	SVM	87.25	86.55	89.74	12.75%

The following table shows the type of tuning to provide the best accuracy in King Abdul-Aziz University's CRM classification.

Classes	Classifier name	Accuracy	Class recall	Class precision	Classification error
Positive, negative, and neutral	NB	95.05	96.69	96.48	4.95%
	SVM	81.80	62.82	55.83	18.20%

XIX shows the classification accuracy, precision, and recall for the NB and SVM classifiers: Crosse-validation=10, sampling type=stratified sampling. The best accuracy was achieved by NB due to the advantages of NB, such as its simplicity, ease of implementation, and combination of efficiency with acceptable accuracy.

TABLE. XIX. ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS

Classes	Classifier name	Accuracy	Class recall	Class precision	Classification error
Positive, negative, and neutral	NB	95.05	96.69	96.48	4.95%
	SVM	81.80	62.82	55.83	18.20%

The following table shows the type of tuning to provide the best accuracy in King Abdul-Aziz University's CRM classification.

Classes	Classifier name	Accuracy	Class recall	Class precision	Classification error
Positive, negative, and neutral	NB	98.33	72.44	72.31	1.67%
	SVM	96.72	65.44	64.52	3.28%

XX shows the classification accuracy, precision, and recall for the NB and SVM classifiers: Crosse-validation=10, sampling type=stratified sampling. The best accuracy was achieved by NB due to the advantages of NB, such as its simplicity, ease of implementation, and combination of efficiency with acceptable accuracy.

TABLE. XX. ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS

Classes	Classifier name	Accuracy	Class recall	Class precision	Classification error
Positive, negative, and neutral	NB	98.33	72.44	72.31	1.67%
	SVM	96.72	65.44	64.52	3.28%

The following table shows the type of tuning to provide the best accuracy in King Abdul-Aziz University's CRM classification. Table XXI shows the classification accuracy, precision, and recall for the NB and SVM classifiers: Crosse-

validation=10, sampling type=stratified sampling. The best accuracy was achieved by NB due to the advantages of NB, such as its simplicity, ease of implementation, and combination of efficiency with acceptable accuracy.

TABLE. XXI. ACCURACY, PRECISION AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS.

Classes	Classifier name	Accuracy	Class recall	Class precision	Classification error
Positive, negative, and neutral	NB	85.13	84.61	85.88	14.87%
	SVM	83.05	81.47	83.12	16.95%

VIII. TUNING OF CRM VIA CEM

This process aims to find a way to support CRM (or Blackboard in this case study). For instance, what exactly do students want or think about experiences? This is especially relevant for students who mostly depend on the Web, such as online distance education students, also proving that the value of social media information can bring better understanding of a student's study situation. Therefore, similar tweets are grouped such that tweets within the same group bear similarity to each other, while tweets in different groups are dissimilar from each other. This will help to understand students' behaviours and find out the most common problems. Moreover, this will give the university the ability to learn about and validate students' data with more support from social media to develop new e-learning criteria to match the inputs from social media.

This case study investigates an alternative solution for supporting CRM by social media inputs on the aspect sentiment level and tuning CRM weights for some aspects of the e-learning criteria that students need, according to their posts on social media. Optimisation of the CRM weights was applied to update some aspects' values in CRM. The results show closely CRM's student labels match CEM's, especially in the negative and neutral classes. Furthermore, they show that optimising CRM's weights can enhance classification accuracy in the Blackboard data and help to understand more students' studying statuses.

A. Experiment in Tuning of CRM via CEM

To be included in this experiment, students should have records in CRM and have completed the survey. Out of the 143 with records on Blackboard, only 79 completed the survey. Table XXII shows the difference between CRM and CEM in this case study before the tuning.

TABLE. XXII. NUMBER OF TWEETS UTILIZED FOR DATA MODEL COMPONENT THREE

Question	CEM/CRM	Positive	Negative	Neutral
Q1	CEM	46	5	28
	CRM	31	23	25
Q2	CEM	29	46	4
	CRM	37	24	18
Q3	CEM	28	26	25
	CRM	36	13	30
Q4	CEM	18	26	35

	CRM	31	20	28
Total CEM		121	103	92
Total CRM		135	80	101

B. Experiment Results and Evaluations

The results were divided into four groups for each question to show the comparison between the CRM and CEM labelling classes.

a) Integration results for Question 1

Fig. 2 shows a comparison between CRM labelling and CEM labelling. Thirty tweets were labelled similarly by CRM and CEM. On the other hand, 48 tweets were labelled dissimilarly by CEM and CRM. In other words, 62% of collected tweets were labelled differently by CRM and CEM, and only 38% were similarly labelled by CEM and CRM.

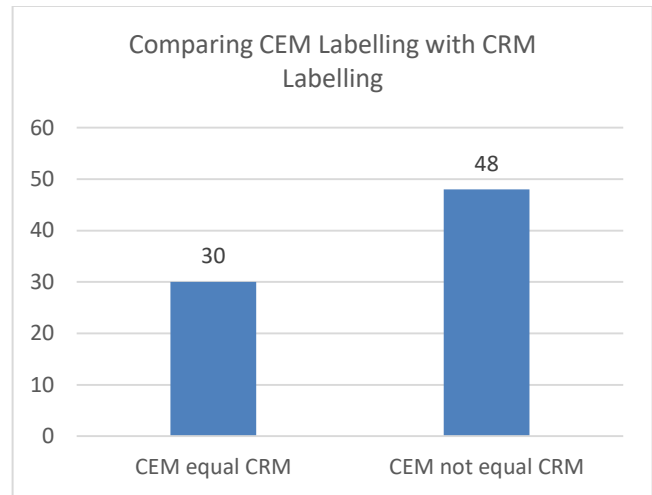


Fig. 2. Comparing CEM labelling with CRM labelling for Question 1.

From Fig. 3, there are 46 positive tweets in CEM, compared to 31 tweets in CRM. In contrast, there are only five negative tweets in CEM, compared to 23 negative tweets in CRM. For the neutral class, there are 28 tweets in CEM and 25 in CRM.

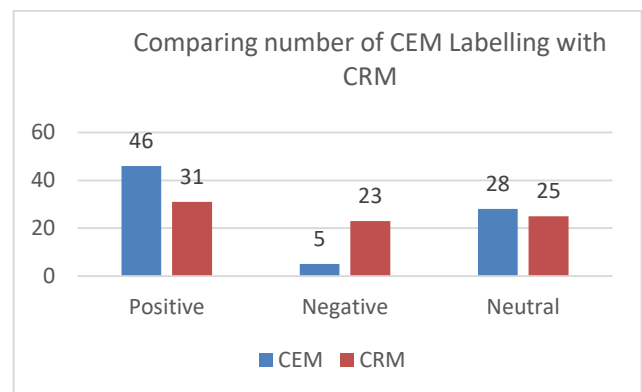


Fig. 3. Comparing CEM labelling class with CRM labelling class for Question 1.

b) Integration results for Question 2

Fig. 4 shows a comparison between CRM labelling and CEM labelling. Twenty-three tweets were labelled similarly between CRM and CEM. On the other hand, 56 tweets were labelled dissimilarly between CEM and CRM. In other words, 71% of collected tweets were labelled differently between CRM and CEM, and only 29% had the same labelling between CEM and CRM.

Fig. 5 shows the number of tweets with similar labelling between the CRM positive class and to the CEM (positive, negative, and neutral) labelling for Question 2. In CEM, 29 tweets were labelled as positive. By contrast, CRM labelled 11 of these tweets as negative, 10 as positive, and eight as neutral.

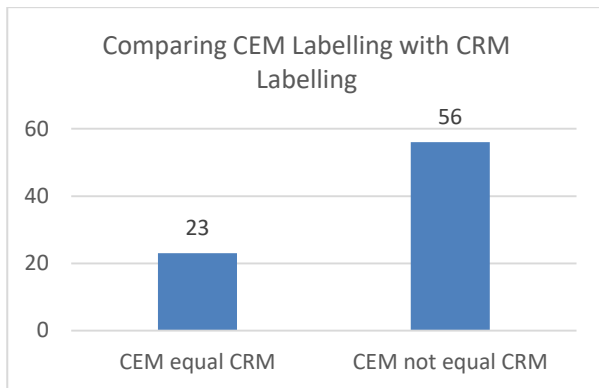


Fig. 4. Comparing CEM labelling with CRM labelling for Question 2.

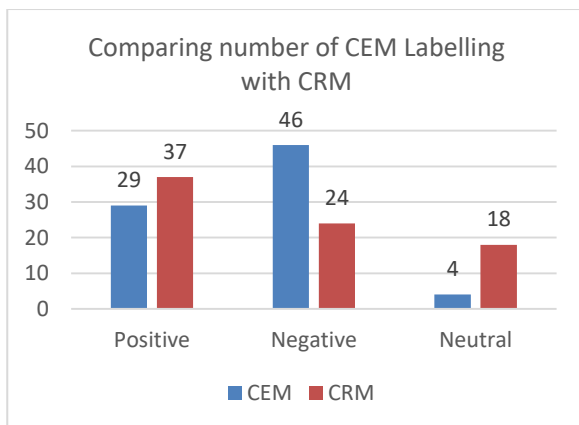


Fig. 5. Comparing CEM labelling with CRM labelling for Question 2.

c) Integration results for Question 3

Fig. 6 shows a comparison between CRM labelling with CEM labelling. Twenty tweets were labelled similarly between CRM and CEM. On the other hand, 59 tweets were labelled dissimilarly between CEM and CRM. In other words, 75% of collected tweets were labelled differently between CRM and CEM, and only 25% had similar labelling.

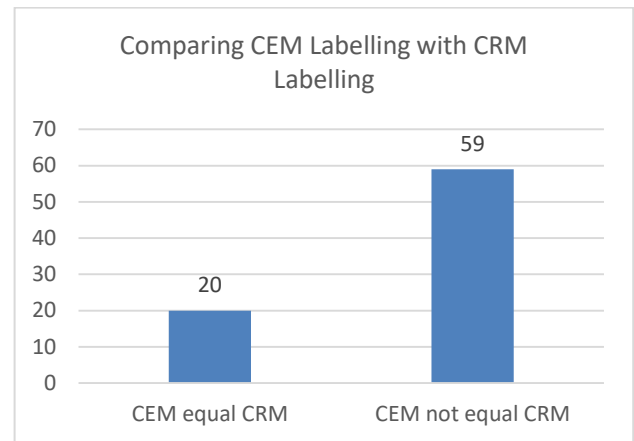


Fig. 6. Comparing CEM labelling with CRM labelling for Question 3

From Fig. 7, there were 28 positive tweets in CEM, compared to 36 tweets in CRM. In contrast, CEM had only 13 negative tweets, compared to 26 in CRM. The neutral class contained 25 tweets in CEM and 30 in CRM, a difference of only five tweets.

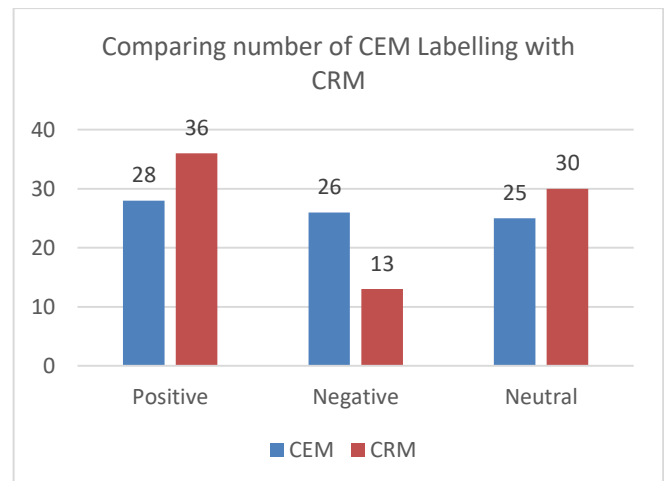


Fig. 7. Comparing CEM labelling with CRM labelling for Question 3.

d) Integration Results for Question 4

Fig. 8 shows a comparison between CRM labelling and CEM labelling. Twenty-four tweets were labelled similarly between CRM and CEM. On the other hand, 55 tweets were labelled dissimilarly between CEM and CRM. In other words, 70% of the collected tweets were labelled differently between CRM and CEM, and only 30% had the same labelling.

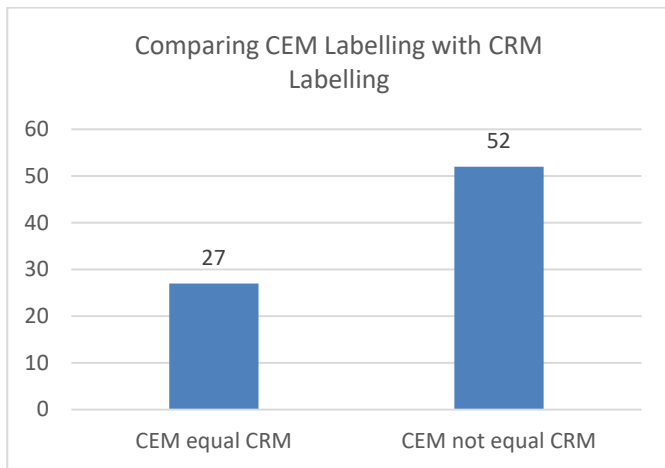


Fig. 8. Comparing CEM labelling with CRM labelling for Question 4.

From Fig. 9, there were 18 positive tweets in CEM, compared to 31 tweets in CRM. In contrast, there were 26 negative tweets in CEM, compared to 20 negative tweets in CRM. For the neutral class, there were 35 tweets in CEM, compared to 28 in CRM.

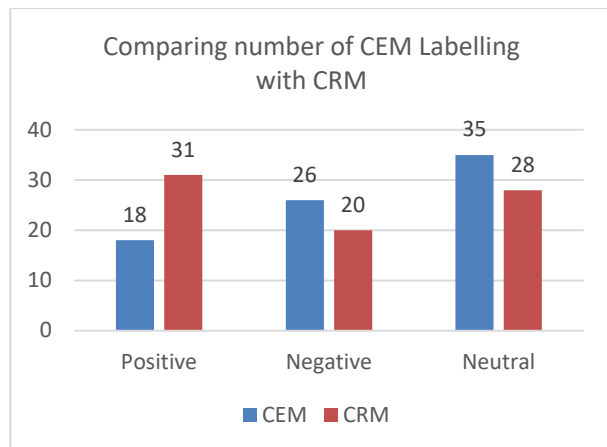


Fig. 9. Comparing CEM labelling with CRM labelling for Question 4.

IX. TUNING THE CRM WEIGHTS

The aim of this experiment was to tune CRM weights for some aspects of e-learning criteria that students need, according to their social media posts. To achieve this aim, first, CRM weights were updated by changing the values of some attributes that reflected the aspect-level opinions in CEM. After that, the criteria were also updated according to the input of the opinions in CEM. Then, the validation was carried out for the new criteria. Last, classification was carried out in order to compare the original criteria with the updated ones. The weight of CRM changed based on the input from social media. For this study, the values of the criteria were changed for aspects of Q1. Changes occurred for students with Blackboard marks between 30 and 40 and student final exam marks in ODUS from 70 to 60. After updating the weight of CRM, another classification was carried out in order to evaluate the accuracy, precision, and recall of all classes for the SVM and NB classifiers. After that, a comparison was carried out to find out the accuracy, precision, and recall of all Classes for the SVM and

NB classifiers before and after CRM weight tuning. Table XXVI shows that the accuracy before the tuning is higher than after the tuning since the matching between the CRM criteria weights and students' opinions on social media are nearly the same for Q1. This indicates that there was no need to update the criteria weights for the aspects in Q1. Table XXIII shows the validation of the classification accuracy, precision, and recall for the NB and SVM classifiers for Question 1: Crosse-validation=10, sampling type=stratified sampling.

TABLE. XXIII. EVALUATION FOR ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS FOR QUESTION 1

Tuning	Classifier name	Accuracy	Class recall	Class precision	Classification error
Before	NB	93.43	93.95	94.12	6.57%
	SVM	87.25	86.55	89.74	12.75%
After	NB	92.58	92.38	92.59	7.42%
	SVM	85.18	83.37	84.70	14.82%

Values of the criteria were changed for aspects in Q2: Changes occurred in student attendance in CRM from seven days to 14 days. Table XXIV shows the comparison between the evaluation of the classification before and after the tuning. The results show that the accuracy after the tuning is higher than before for the aspects in Q2.

TABLE. XXIV. EVALUATION FOR ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS FOR QUESTION 2

Tuning	Classifier name	Accuracy	Class recall	Class precision	Classification error
Before	NB	95.05	96.69	96.48	4.95%
	SVM	81.80	62.82	55.83	18.20%
After	NB	95.85	96.09	95.76	4.15%
	SVM	84.70	65.96	57.98	15.30%

Values of the criteria were changed for aspects in Q3: Changes occurred in evaluating the total number of posts, including discussions and tests. Table XXV shows the comparison between the evaluation of the classification before and after the tuning. The results show that the accuracy after the tuning is almost the same.

TABLE. XXV. EVALUATION FOR ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS FOR QUESTION 3

Tuning	Classifier name	Accur acy	Class recall	Class precision	Classificati on error
Before	NB	98.33	72.44	72.31	1.67%
	SVM	96.72	65.44	64.52	3.28%
After	NB	98.35	72.44	72.37	1.65%
	SVM	95.92	64.96	63.92	4.08%

Q4 combines the aspects in Q1, Q2, and Q3. Table XXVI shows the comparison between the evaluation of the classification before and after the tuning. The results show that the accuracy after the tuning is higher than before. This indicates that changing some aspects of the criteria can help to enhance the classification accuracy in the CRM data.

TABLE. XXVI. EVALUATION FOR ACCURACY, PRECISION, AND RECALL WITH ALL CLASSES FOR SVM AND NB CLASSIFIERS FOR QUESTION 4

Tuning	Classifier name	Accuracy	Class recall	Class precision	Classification error
Before	NB	85.13	84.61	85.88	14.87%
	SVM	83.05	81.47	83.12	16.95%
After	NB	92.58	92.38	92.59	7.42%
	SVM	85.18	83.37	84.70	14.82%

To sum up King Abdul-Aziz University's validation CRM experiments, the best accuracy was achieved by NB due to the advantages of NB such as its simplicity, ease of implementation, and combination of efficiency and acceptable accuracy.

Fig. 10 shows a comparison between the CEM labelling and the distance education criteria with suggestion one for Question 1. The number of positive tweets increases, as does the number of negative tweets, while the number of neutral tweets sharply decreases.

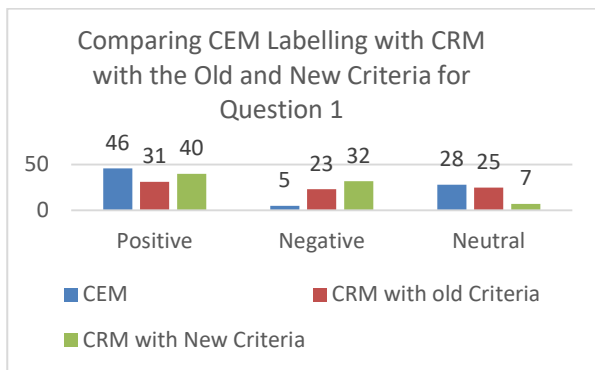


Fig. 10. Comparing CEM labelling with CRM with the old and new criteria for Question 1.

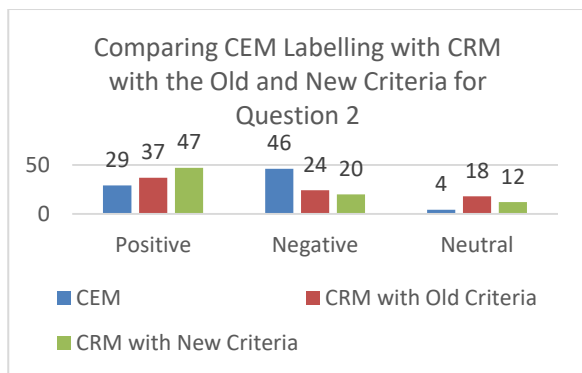


Fig. 11. Comparing CEM labelling with CRM with the old and new criteria for Question 2.

Fig. 11 shows a comparison between the CEM labelling and the distance education criteria with suggestion one for Question 1. The number of positive tweets increases, and the number of negative tweets decreases, as does the number of neutral tweets.

Fig. 12 shows a comparison between the CEM labelling and the distance education criteria with suggestion one for Question 1. The number of positive tweets approximately doubles, and the number of negative tweets approximately halves. In addition, the number of neutral tweets drops to zero.

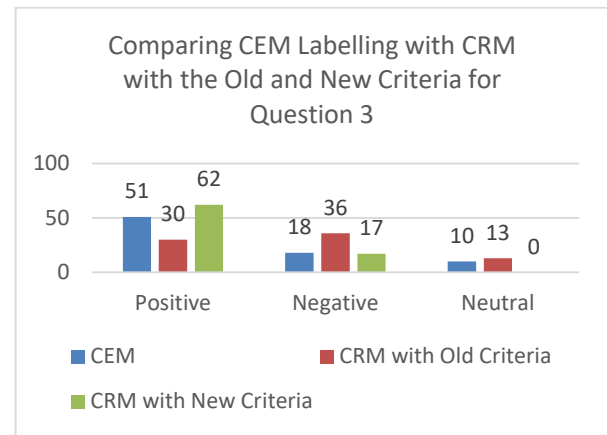


Fig. 12. Comparing CEM labelling with CRM with the old and new criteria for Question 3.

Fig. 13 shows a comparison between the CEM labelling and the distance education criteria with suggestion one for Question 1. The number of positive and neutral tweets increases, and the number of negative tweets decreases.

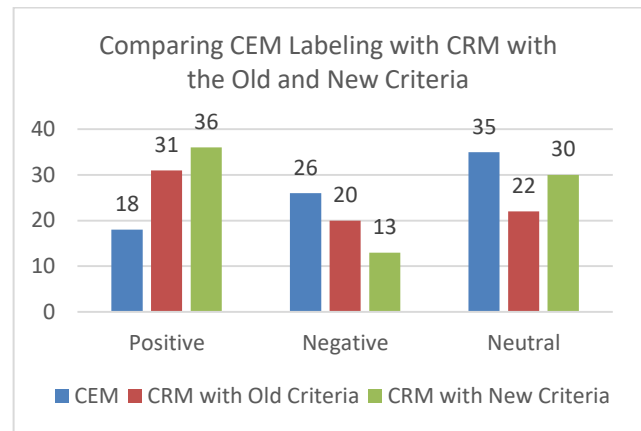


Fig. 13. Comparing CEM labelling with CRM with the old and new criteria for Question 4.

In conclusion, the result of this experiment shows that social media can support CRM with more details. The main aim of this study was achieved by showing the gap between the criteria and students' needs. The validation result demonstrates the gap between the students' perspectives and the criteria. This might help universities to adapt the distance education criteria in a way that helps the students to deal with the university. The validation results confirm that tuning the criteria will help students. However, the main reasons for the difference of the classification results with the given examples above is domain diversity, as well as the way that data were pre-processed, the size of the dataset, and of course the approaches that were used in each article. Moreover, In terms of the accuracy of the last survey, which studied recent work in Arabic sentiment analysis, SVM was applied successfully in several sentiment analysis tasks [27], [28].

X. CONCLUSION

This study carried out to investigate whether social media could help experts to understand users' perspectives and could

support academics' knowledge about their students. Therefore, a framework was proposed for integrating CRM with CEM on the aspect level. The framework consists of three components: sentiment analysis for CEM, classification using CRM Blackboard data, and tuning of CRM via CEM, which integrates both results to study the level of matching between both resources' information, namely, social media and Blackboard data. In other words, there is good consistency between the CRM structure and students' opinions on the aspect level. This takes the study further by investigating King Abdul-Aziz University in e-learning and distance criteria, which brings the similarity between CEM and CRM opinions closer. Moreover, the final stage of this experiment shows an interesting result after changing the e-learning criteria according to the necessary requirements of input on social media requested in students' feedback comments through the Twitter platform. The results show that changing some aspects of the e-learning criteria that were required by students in their social media posts can help to enhance the classification accuracy in the Blackboard data and to understand more students' studying statuses. Furthermore, the results show matching between students' opinions in CRM and CEM, especially in the negative and neutral classes.

REFERENCES

- [1] Tripathi, G. and S. Naganna, Feature selection and classification approach for sentiment analysis. *Machine Learning and Applications: An International Journal*, 2015. 2(2): p. 1-16.
- [2] Dehkharghani, R. and C. Yilmaz. Automatically identifying a software product's quality attributes through sentiment analysis of tweets. in *Natural Language Analysis in Software Engineering (NaturaLiSE)*, 2013 1st International Workshop on. 2013. IEEE.
- [3] Al-Rubaiee, H., R. Qiu, and D. Li. Identifying Mubasher software products through sentiment analysis of Arabic tweets. in *Industrial Informatics and Computer Systems (CIICS)*, 2016 International Conference on. 2016. IEEE.
- [4] Bross, J. and H. Ehrig. Generating a context-aware sentiment lexicon for aspect-based product review mining. in *Web Intelligence and Intelligent Agent Technology (WI-IAT)*, 2010 IEEE/WIC/ACM International Conference on. 2010. IEEE.
- [5] Indhuja, K. and R.P. Reghu. Fuzzy logic based sentiment analysis of product review documents. in *Computational Systems and Communications (ICCSC)*, 2014 First International Conference on. 2014. IEEE.
- [6] Wang, W. Sentiment analysis of online product reviews with Semi-supervised topic sentiment mixture model. in *Fuzzy Systems and Knowledge Discovery (FSKD)*, 2010 Seventh International Conference on. 2010. IEEE.
- [7] Zhang, Z. and B. Varadarajan. Utility scoring of product reviews. in *Proceedings of the 15th ACM international conference on Information and knowledge management*. 2006. ACM.
- [8] Bollen, J., H. Mao, and X. Zeng. Twitter mood predicts the stock market. *Journal of Computational Science*, 2011. 2(1): p. 1-8.
- [9] Hamed, A.-R., R. Qiu, and D. Li. Analysis of the relationship between Saudi twitter posts and the Saudi stock market. in *Intelligent Computing and Information Systems (ICICIS)*, 2015 IEEE Seventh International Conference on. 2015. IEEE.
- [10] Martin, V., E. Bruno, and E. Murisasco, Predicting the French Stock Market Using Social Media Analysis. *International Journal of Virtual Communities and Social Networking (IJVCSN)*, 2015. 7(2): p. 70-84.
- [11] Şimşek, M.U. and S. Özdemir. Analysis of the relation between Turkish twitter messages and stock market index. in *Application of Information and Communication Technologies (AICT)*, 2012 6th International Conference on. 2012. IEEE.
- [12] Khatri, S.K., H. Singhal, and P. Johri. Sentiment analysis to predict Bombay stock exchange using artificial neural network. in *Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)*, 2014 3rd International Conference on. 2014. IEEE.
- [13] Gao, Y., et al. Sentiment classification for stock news. in *Pervasive Computing and Applications (ICPCA)*, 2010 5th International Conference on. 2010. IEEE.
- [14] Zhang, W. and S. Skiena. Trading Strategies to Exploit Blog and News Sentiment. in *Icwsn*. 2010.
- [15] Hamed AL-Rubaiee, R.Q., Khalid Alomar and Dayou Li, Sentiment Analysis of Arabic Tweets in e-Learning. *Journal of Computer Science*. 12(11): p. 553-563.
- [16] Tian, F., et al. Can e-Learner's emotion be recognized from interactive Chinese texts? in *Computer Supported Cooperative Work in Design*, 2009. CSCWD 2009. 13th International Conference on. 2009. IEEE.
- [17] Wang, P., M. Zuo, and L. Diao. Design of the monitoring and analysis system of the Internet Education Public Sentiment. in *Future Computer and Communication (ICFCC)*, 2010 2nd International Conference on. 2010. IEEE.
- [18] Fu, G. and X. Wang. Chinese sentence-level sentiment classification based on fuzzy sets. in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*. 2010. Association for Computational Linguistics.
- [19] Donovan, J., C.E. Mader, and J. Shinsky. Constructive Student Feedback: Online vs. Traditional Course Evaluations. *Journal of Interactive Online Learning*, 2010. 9(3): p. 283-296.
- [20] Mohan, S., E. Choi, and D. Min. Conceptual modeling of enterprise application system using social networking and Web 2.0 "social CRM system". in *Convergence and Hybrid Information Technology*, 2008. ICHIT'08. International Conference on. 2008. IEEE.
- [21] Biere, M., *Business intelligence for the enterprise*. 2003: Prentice Hall Professional.
- [22] Spangler, S., et al., COBRA-Mining web for corporate brand and reputation analysis. *Web Intelligence and Agent Systems: An International Journal*, 2009. 7(3): p. 243-254.
- [23] Sukumaran, S. and A. Sureka, Integrating structured and unstructured data using text tagging and annotation. *Business Intelligence Journal*, 2006. 11(2): p. 8.
- [24] Ajmera, J., et al. A CRM system for social media: challenges and experiences. in *Proceedings of the 22nd international conference on World Wide Web*. 2013. ACM.
- [25] Yaakub, M.R., Y. Li, and J. Zhang. Integration of sentiment analysis into customer relational model: the importance of feature ontology and synonym. *Procedia Technology*, 2013. 11: p. 495-501.
- [26] Hu, M. and B. Liu, Mining and summarizing customer reviews, in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 2004, ACM: Seattle, WA, USA. p. 168-177.
- [27] Boudad, N., et al., Sentiment analysis in Arabic: A review of the literature. *Ain Shams Engineering Journal*, 2017.
- [28] Kaseb, G.S. and M.F. Ahmed, Arabic Sentiment Analysis approaches: An analytical survey.